APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
10/767,151	01/29/2004	Parvathi Chundi	10990670-2	4632

			EXAMINER	
7590	05/16/2006		BLACKWELL, JAMES H	

HEWLETT-PACKARD COMPANY
Intellectual Property Administration
P.O. Box 272400
Fort Collins, CO 80527-2400

ART UNIT	PAPER NUMBER
2176	

DATE MAILED: 05/16/2006

Please find below and/or attached an Office communication concerning this application or proceeding.

	Application No.	Applicant(s)
	10/767,151	CHUNDI ET AL.
Office Action Summary	Examiner	Art Unit	
	James H. Blackwell	2176	

-- *The MAILING DATE of this communication appears on the cover sheet with the correspondence address* --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE <u>3</u> MONTH(S) OR THIRTY (30) DAYS, WHICHEVER IS LONGER, FROM THE MAILING DATE OF THIS COMMUNICATION.
- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

1)☒ Responsive to communication(s) filed on <u>*29 January 2004*</u>.

2a)☐ This action is **FINAL**.   2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

4)☒ Claim(s) *1-35* is/are pending in the application.

    4a) Of the above claim(s) *1-16* is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *17-35* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

9)☐ The specification is objected to by the Examiner.

10)☒ The drawing(s) filed on <u>*29 January 2004*</u> is/are: a)☒ accepted or b)☐ objected to by the Examiner.

    Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

    Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

    a)☐ All   b)☐ Some * c)☐ None of:

      1.☐ Certified copies of the priority documents have been received.

      2.☐ Certified copies of the priority documents have been received in Application No. _____.

      3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

    * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

1) ☒ Notice of References Cited (PTO-892)

2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)

3) ☒ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date <u>*01/29/04*</u>.

4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____ .

5) ☐ Notice of Informal Patent Application (PTO-152)

6) ☐ Other: _____ .

DETAILED ACTION

1.   This Office Action is in response to an original application filed 01/29/2004 with

apriority date of **03/22/2000**.

2.   Claims 17-35 are pending.

3.   Claims 1-16 were cancelled by Applicant by pre-amendment.

4.   Claims 17, 23, 28, 31, and 35 are independent claims.

Claim Objections

5.   Claims 29 and 30 are objected to because of the following informalities: Both of

these claims refer to Claim 9, which was cancelled. Appropriate correction is required.

Claim Rejections - 35 USC § 103

6.   The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all

obviousness rejections set forth in this Office action:

> (a) A patent may not be obtained though the invention is not identically disclosed or described as set
> forth in section 102 of this title, if the differences between the subject matter sought to be patented and
> the prior art are such that the subject matter as a whole would have been obvious at the time the
> invention was made to a person having ordinary skill in the art to which said subject matter pertains.
> Patentability shall not be negatived by the manner in which the invention was made.

7.   Claims 17 and 18 are rejected under 35 U.S.C. 103(a) as being unpatentable

over Cooley et al. (hereinafter Cooley, "WebSIFT: The Web Site Information Filter

System", Copyright 06/13/1999).

In regard to independent Claim 17, and similarly dependent Claim 18,

Cooley teaches the method of Web Usage Mining. This method involves the application

of data mining techniques, including clustering, to large Web data repositories using

server logs and the HTML files that make up the web site (*documents*), in order to

produce clusters from which usage patterns can be extracted (see Abstract, Fig. 2).

Claim 17 recites *a method for generating clusters based on a combination of web*

session logs (i.e., server logs) and the HTML files that make up the site in order to

produce clusters that incorporate a users' perspective.

Cooley also discusses the notion of a distance between two web documents and

how that relates to the similarity between them (Sec. 3) as read in Claim 18.

Cooley does not explicitly teach about log-based or content-based clustering or

how to make the Euclidean Distance between documents the same. However, the log-

based clustering method described in Claim 17 contains steps that use content-based

clustering. Furthermore, content-based clustering in Claim 17 would have been obvious

to one of ordinary skill in the art at the time of invention because the steps involved in

creating the clusters of Claim 17 were well known in the art of clustering at the time of

invention as is the notion of a distance between documents as expressed in Claim 18

(e.g., see Jain et al., Pg. 267, 2nd Paragraph). One of ordinary skill in the art at the time

of invention would have been motivated to follow Cooley's general approach to

clustering and computing distances (similarities) between documents because it follows

similar steps to what one would have generally done in preparing documents to be

clustered especially in the sense of content-based clustering as described in Claims 17

and 18.

In regard to dependent Claims 19-22, Cooley fails to explicitly teach that *each*

session log comprises a query used to retrieve documents or *a number of documents*

found to satisfy a query or *a list of documents opened by a user* or *a length of time that a document was opened.* However, the structure of Web session logs (e.g., Common Log Format, Extended Log Format) was well known to one of ordinary skill in the art at the time of invention and therefore obvious (see W3C Common Log Format, Extended Log Format).

8.   Claims 23-27 are rejected under 35 U.S.C. 103(a) as being unpatentable over Cutting et al. (hereinafter Cutting, "Scatter/Gather: A Cluster-based Approach to Browsing Large Document Collections", Copyright circa 1992 ACM).

In regard to independent Claim 23, and similarly dependent Claims 23-27, Cutting discusses the notion of a sparse vector and how it is constructed from the values of document parameters that are similar (Sec. 3, Par. 2). Furthermore, it states that the vectors can be represented by Boolean one's and zero's. This would make any hybrid matrix constructed from such vector elements easier to deal with once a clustering algorithm was applied.

Cutting does not teach the specific method of constructing the hybrid matrix or the specific clustering algorithm as read in Claims 23-27. However, it would have been obvious to one of ordinary skill in the art at the time of invention to construct vectors based on a document's clustering parameters and to "normalize" those vectors to Boolean values when constructing a matrix because this was a common procedure to follow when one prepared to apply a clustering algorithm to a set of document data (e.g., see Jain et al., Abstract; Pg. 268 discusses normalizing).

9.   Claims 28 and 35 are rejected under 35 U.S.C. 103(a) as being unpatentable

over Cooley in view of Pitkow et al. (hereinafter Pitkow, U.S. Patent No. 6,457,028 filed

09/29/1999, issued 09/24/2002).

In regard to independent Claims 28 and 35, Claims 28, and 35 reflect the

method for clustering documents, including generating clusters with user perspective,

as claimed in Claim 17, and is rejected along the same rationale. In addition, Cooley

does not teach a processor and external storage. However, Pitkow teaches a processor

and External Storage Device (Col. 12, lines 12-13; Fig. 10) and Internal Memory which

is a combination of both Random Access (RAM) and Read-only (ROM) memory. It

would have been obvious to one of ordinary skill in the art at the time of invention to

combine the teachings of Cooley and Pitkow as both inventions relate to document

clustering. Adding the teaching of Pitkow allows for storage of clustering data.

10.   Claims 29 and 30 are rejected under 35 U.S.C. 103(a) as being unpatentable

over Cooley in view of Pitkow, and in further view of Cutting.

In regard to dependent Claims 29 and 30, Cooley fails to teach documents

stored in storage, as claimed in Claim 29. However, Pitkow teaches an External Storage

Device, Internal Memory connected to a processor (Col. 12, Fig. 10). It would have

been obvious to one of ordinary skill in the art at the time of invention to combine the

teachings of <u>Cooley</u> and <u>Pitkow</u> as both inventions relate to document clustering.

Adding the teaching of <u>Pitkow</u> allows for storage of clustering data.

    <u>Pitkow</u> does not teach a hybrid matrix comprising the log-based document cluster vectors and individual document vectors, as claimed in Claim 30. However, <u>Cutting</u> discusses the notion of a sparse vector and how it is constructed from the values of document parameters that are similar (Sec. 3, Par. 2). It would have been obvious to one of ordinary skill in the art at the time of invention to combine the teachings of <u>Cooley,</u> <u>Pitkow</u>, and <u>Cutting</u> as all three inventions relate to document clustering. Adding the teaching of <u>Cutting</u> allows for storage of clustering data in an efficient manner.

11.   Claims 31-34 are rejected under 35 U.S.C. 103(a) as being unpatentable over

Pitkow in view of Cooley.

In regard to independent Claim 31, Claim 31 reflects the method for clustering

documents, including generating clusters with user perspective as claimed in Claim 17,

and is rejected along the same rationale. In addition, Pitkow teaches a processor and

External ClaimStorage Device (Col. 12, lines 12-13; Fig. 10), which can include fixed or

removable magnetic or optical disk drive (Col. 12, lines 38-39; Fig. 10), and an Internal

Memory which is a combination of both Random Access (RAM) and Read-only (ROM)

memory (Col. 12, lines 13-16; Fig. 10).

Cooley teaches session logs and documents and the general notion of clustering

documents and logs together. Pitkow does not teach a document clustering module

having a plurality of instructions, which when executed by the processor, performs log-

based clustering on the session logs to generate session clusters, converts the session

clusters into a form suitable for content-based clusters, performs content-based

clustering to generate document clusters with users' perspective.

In regard to dependent Claim 34, Cooley does not teach the specific method of

combining the session logs and the documents to perform the clustering described in

Claim 31. However, one of ordinary skill in the art at the time of invention would have

been motivated to combine Pitkow and Cooley because storing a clustering program on

an external storage device allows one to later retrieve and execute it on a processor.

Conclusion

12.   Any inquiry concerning this communication or earlier communications from the examiner should be directed to James H. Blackwell whose telephone number is 571-272-4089. The examiner can normally be reached on Mon-Fri.

13.   If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Heather R. Herndon can be reached on 571-272-4136. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

14.   Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

James H. Blackwell
04/25/2006

WILLIAM BASHORE
PRIMARY EXAMINER
5/1/2006